Original Article

After continents divide: comparative phylogeography of reef fishes from the Red Sea and Indian Ocean

Joseph D. DiBattista[1*], Michael L. Berumen[2,3], Michelle R. Gaither[4], Luiz A. Rocha[4], Jeff A. Eble[5], J. Howard Choat[6], Matthew T. Craig[7], Derek J. Skillings[1] and Brian W. Bowen[1]

[1]Hawai'i Institute of Marine Biology, Kāne'ohe, HI 96744, USA, [2]Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, [3] Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543 USA, [4]Section of Ichthyology, California Academy of Sciences, San Francisco, CA 94118, USA, [5]Department of Biology, University of West Florida, Pensacola, FL 32514, USA, [6]School of Marine and Tropical Biology, James Cook University, Townsville, QLD 4811, Australia, [7]Department of Marine Sciences and Environmental Studies, University of San Diego, San Diego, CA 92110, USA

*Correspondence: Joseph D. DiBattista, Hawai'i Institute of Marine Biology, P.O. Box 1346, Kāne'ohe, HI 96744, USA.

E-mail: joseph99@hawaii.edu

Running header: Phylogeography of Red Sea reef fishes

ABSTRACT

Aim The Red Sea is a biodiversity hotspot characterized by unique marine fauna and high endemism. This sea began forming approximately 24 million years ago with the separation of the African and Arabian plates, and has been characterized by periods of desiccation, hypersalinity and intermittent connection to the Indian

Ocean. We aim to evaluate the impact of these events on the genetic architecture of the Red Sea reef fish fauna.

Location Red Sea and Western Indian Ocean

Methods We surveyed seven reef fish species from the Red Sea and adjacent Indian Ocean using mitochondrial DNA cytochrome-c oxidase subunit I and cytochrome *b* sequences. To assess genetic variation and evolutionary connectivity within and between these regions, we estimated haplotype diversity and nucleotide diversity, reconstructed phylogenetic relationships among haplotypes and estimated gene flow and time of population separation using Bayesian coalescent-based methodology.

Results Our analyses revealed a range of scenarios from shallow population structure to diagnostic differences that indicate evolutionary partitions and possible cryptic species. Conventional molecular clocks and coalescence analyses indicated time frames for divergence between these bodies of water ranging from 830,000 years to contemporary exchange or range expansion. Colonization routes were bidirectional, with some species moving from the Indian Ocean to the Red Sea compared with expansion out of the Red Sea for other species.

Main conclusions We conclude that: (1) at least some Red Sea reef fauna survived multiple salinity crises; (2) endemism is higher in the Red Sea than previously reported; and (3) the Red Sea is an evolutionary incubator, occasionally contributing species to the adjacent Indian Ocean. The latter two conclusions – elevated endemism and species export – indicate a need for enhanced conservation priorities for the Red Sea.

Keywords

Coalescent, cryptic speciation, dispersal, genealogical concordance, gene flow, mitochondrial DNA, vicariance.

INTRODUCTION

The Red Sea is a deep (maximum depth: 2920 m) and narrow (maximum width: 350 km) body of water extending 2270 km from 30° N in the Gulf of Suez to 13° N in the Gulf of Aden. Although the sea began forming approximately 24 million years ago (Ma) (i.e. mid-Oligocene period) by the separation of the African and Arabian plates, the ocean environment that supports coral reefs originated in the

78 early Pleistocene (*c.* 5 Ma; Siddall *et al.*, 2003; Boswell *et al.*, 2005). The Red
79 Sea, which now experiences minimal freshwater inflow and high rates of
80 evaporation, is characterized by pronounced north to south gradients in salinity
81 (42 ppt to 37 ppt), sea surface temperature (winter, 20 °C to 28 °C; summer, 26
82 °C to 32 °C), and nutrient concentration (low to high) (Raitsos *et al.*, 2011; Ngugi
83 *et al.*, 2012). Oceanographic current patterns and climate in the southern Red Sea
84 (and Gulf of Aden) are heavily influenced by the northern Indian Ocean monsoon
85 system (Smeed, 1997, 2004; Biton *et al.*, 2010), which reverses wind circulation
86 patterns in the boreal summer (Southwest Monsoon, April to October) compared
87 to those in the winter (Northeast Monsoon, December to March; see Fig. 1).
88 Despite its peripheral location relative to the Indo-Pacific, the Red Sea is
89 characterized by high biodiversity, including approximately 300 reef-building
90 corals (mostly species of *Acropora* and *Porites*; Sheppard & Sheppard, 1991;
91 Riegl & Velimirov, 1994) and 1078 fish species (Golani & Bogorodsky, 2010),
92 which represent key resources for coastal communities. Species richness appears
93 to be highest in the central Red Sea, with marked decreases in species abundance
94 and changes in species composition away from this area (DeVantier *et al.*, 2000).
95 The Red Sea also harbours one of the highest degrees of endemism for marine
96 organisms, making up 14% of fishes (Randall, 1994), 33% of crustaceans, 15% of
97 echinoderms, and up to 25% of corals (Cox & Moore, 2000). Endemism can be

even higher for some taxa, reaching 50% in the butterflyfishes (Roberts *et al.*, 1992).

Although the evolutionary processes driving high rates of endemism are unclear, the narrow (18 km) and shallow (137 m) Straight of Bab al Mandab, the only connection with the Indian Ocean, is likely to have played a key role (Klausewitz, 1989). The Red Sea was repeatedly isolated during Pleistocene glacial cycles when the sea level lowered as much as 120 m; whether this was achieved through physical isolation or the restriction of oceanic flow associated with elevated salinities and temperatures remains contentious (Siddall *et al.*, 2003; Bailey, 2009). Moreover, cold-water upwelling off Somalia, which precludes reef formation on the northeast African and southern Arabian coasts, is likely to reinforce this isolation (Smeed, 1997; Kemp, 1998, 2000). Some authors believe that Bab al Mandab no longer acts as a physical barrier to dispersal but that an ecological barrier lies within the Red Sea (Ormond & Edwards, 1987). Roberts *et al.* (1992) suggested that a turbid-water region south of 20° N in the Red Sea may limit larval dispersal, a hypothesis which is supported by the presence of a number of species in the northern/central Red Sea and the Gulf of Aden (just outside the Red Sea) that are absent from the southern Red Sea.

Despite being acknowledged as a biodiversity hotspot for coral reef fishes based on research dating back more than 200 years (e.g. Forsskål 1775), little work has been conducted in the Red Sea using modern genetic techniques.

Studies in this region tend to focus on the biogeography and community structure of the more iconic (and endemic) shore fish fauna (e.g. family Chaetodontidae; Righton *et al.*, 1996). The majority of genetic studies on reef fishes have been restricted to the Gulf of Aqaba and northern Red Sea (Hassan *et al.*, 2003; Kochzius & Blohm, 2005; but see Froukh & Kochzius, 2007), and few of these considered the connections between widespread taxa from other biogeographical provinces, particularly the Indian Ocean (Froukh & Kochzius, 2008).

Peripheral reef habitats such as the Red Sea, which forms the north-western most extension of the Indian Ocean, are typically considered to be biodiversity sinks that receive species from elsewhere but rarely export them (Briggs, 1999). The accepted paradigm is therefore that biogeographical sinks are 'evolutionary graveyards' that do not contribute to biodiversity at neighbouring sites. Recent research on reef fish and invertebrates, however, demonstrate that peripheral regions, such as the Hawaiian Archipelago and the Marquesas Islands, may act as both a sink and a source, contributing unique genetic lineages to other regions of the Indo-Pacific (Gaither *et al.*, 2010, 2011; DiBattista *et al.*, 2011; Eble *et al.*, 2011; Skillings *et al.*, 2011).

Our first aim is to assess connections between fauna in the Red Sea and the adjacent Western Indian Ocean (WIO) using a molecular genetic approach. The WIO forms a biogeographical subdivision of the tropical Indo-Pacific stretching from East Africa to the Chagos Ridge in the centre of the Indian Ocean

(Sheppard, 2000; Briggs & Bowen, 2012). Phylogeographical inferences are strengthened by congruence among multiple species and genes, and so our study considers seven species of reef fish with widespread distributions using two mitochondrial DNA (mtDNA) markers.

Our second aim is to assess whether sea level changes have influenced extant biodiversity by estimating migration rates and divergence times of reef fishes in the Red Sea and WIO. Such analyses will allow us to discriminate between the following scenarios: (1) Red Sea populations represent long isolated relicts derived from the WIO, which implies gene flow was absent over the last five million years; (2) Red Sea populations have been isolated from the WIO over evolutionary intervals but with recurrent gene flow; or (3) Red Sea populations are the result of recent colonization from the WIO, since the Last Glacial Maximum approximately 20,000 years ago (Siddall *et al.*, 2003; Bailey, 2009). This dataset provides an unprecedented opportunity to assess the relationships between two Indian Ocean biogeographical provinces, and thereby illuminate evolutionary processes that are the wellspring of Red Sea biodiversity.

MATERIALS AND METHODS

Sample collection

Reef fish were collected while SCUBA diving or snorkelling at depths of 1–40 m between 2002 and 2011 (Fig. 1, Table 1). Seven reef fish species were targeted: the brown surgeonfish, *Acanthurus nigrofuscus* (Forsskål, 1775); the peacock hind, *Cephalopholis argus* (Schneider, 1801); the threadfin butterflyfish, *Chaetodon auriga* (Forsskål, 1775); the checkerboard wrasse, *Halichoeres hortulanus* (Lacepède, 1801); the bluestripe snapper, *Lutjanus kasmira* (Forsskål, 1775); the Sammara squirrelfish, *Neoniphon sammara* (Forsskål, 1775); and the regal angelfish, *Pygoplites diacanthus* (Boddaert, 1772). These species were chosen because they have wide Indo-Pacific distributions, are abundant, represent a diversity of taxonomic families, and can be unequivocally identified in the field. Each species was sampled at two locations (Thuwal and Al Lith) off the coast of the Kingdom of Saudi Arabia (KSA) in the central Red Sea, and at one to three sites in the WIO (oceanic sites: Diego Garcia in the Chagos Archipelago and the Republic of Seychelles; coastal sites: Sodwana Bay, South Africa and Al Hallaniyat, Sultanate of Oman). Because some of the collections were opportunistic, not every species could be sampled at every location (Fig. 1, Table 1).

Mitochondrial DNA sequencing

Tissue samples were preserved in salt-saturated DMSO (Seutin *et al.*, 1991). Total genomic DNA was extracted using the 'HotSHOT' protocol (Meeker *et al.*, 2007) and subsequently stored at −20 °C. Fragments of mtDNA from the cytochrome-c oxidase subunit-I (COI) and cytochrome *b* (cyt*b*) genes were amplified using either previously published primers or modified primers designed for individual species (Table 1). These two markers were chosen because they: (1) are easy to amplify in most fish; (2) are generally variable at the population level; (3) facilitate comparisons with published sequences; and (4) have had molecular clock rates estimated based on reef fishes (Bowen *et al.*, 2001; Lessios, 2008; Reece *et al.*, 2010). Polymerase chain reaction (PCR) was carried out for all species as described in DiBattista *et al.* (2012a), with optimal annealing temperatures listed in Table 1. All samples were sequenced in the forward direction (and reverse direction for unique or questionable haplotypes) with fluorescently labelled dye terminators (BigDye version 3.1, Applied Biosystems , Foster City, CA, USA) and analysed using an ABI 3130xl Genetic Analyzer (Applied Biosystems). The sequences were aligned, edited, and trimmed to a uniform length using GENEIOUS PRO 4.8.4 (Drummond *et al.*, 2009); unique mtDNA haplotypes were deposited in GenBank (accession numbers: KC187734-KC188056.

Genetic diversity and population structure

ARLEQUIN 3.5 (Excoffier *et al.*, 2005) was used to calculate haplotype (*h*) and nucleotide diversity (π), and to test for population structure among sampling sites for each species and molecular marker (i.e. 14 total datasets). These analyses were repeated with all Red Sea or WIO samples pooled into two separate regions. Despite the difference in the geographical scale of sampling (Red Sea sites, *c*. 300 km; WIO sites, ~ 1000s of km), preliminary work suggests that Red Sea haplotypes at Thuwal and Al Lith are consistent with those sampled up to 1200 km north (J.D.D., unpub. data), indicating unbiased estimates of genetic diversity within our study range. Because JMODELTEST 0.1.1 (Posada, 2008) converged on different models of nucleotide sequence evolution among datasets, we calculated global and pairwise Φ_{ST} values based on a HKY model of mutation (Hasegawa *et al.*, 1985). We also ran conventional frequency based F_{ST}, but the absolute values changed little and relative values did not change at all; we have therefore elected to report pairwise Φ_{ST}. Global Φ_{ST} was estimated using analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992); deviations from null distributions were tested with non-parametric permutation procedures (n = 99,999). We controlled for false discovery rate with the method of Narum (2006), and negative pairwise Φ_{ST} values were converted to zeros. To facilitate comparisons among species, an additional diversity measure – Jost's *D* (Jost, 2008) – was estimated using SPADE 1.0 (Chao *et al.*, 2008).

The evolutionary relationship among COI or cyt *b* haplotypes was resolved for each species with unrooted networks constructed with the program NETWORK 4.5.1.0 (www.fluxus-engineering.com/network_terms.htm) using a median-joining algorithm and default settings (as per Bandelt *et al.*, 1999).

IMA2 analysis

We calculated the effects of time and gene flow on genetic divergence between populations using Bayesian coalescent-based estimation with IMA2 8.26.11 (Hey & Nielsen, 2007; Hey, 2010). Using *F*-statistics we determined that samples within regions were not significantly different for all seven species after correction for multiple comparisons. We therefore pooled the Red Sea sites together and the WIO sites together, for comparisons between regions for each species and molecular marker.

The isolation-with-migration model implemented in IMA2 assumes that two populations of effective size N_1 and N_2 diverged from an ancestral population (of effective size N_a) at time t, and then exchanged migrants at rates m_1 and m_2. We therefore estimated the time since initial separation or last major colonization event (t), effective population size (N_e), and the proportion of migrants arriving into a population per generation (m); all demographic parameters were scaled by mutation rate.

Mutation rates calibrated in other reef fish based on the closure of the Isthmus of Panama range from 1% to 2 % per million years for COI and cyt *b* (Bowen *et al.*, 2001; Lessios, 2008; Reece *et al.*, 2010). We used a conservative estimate of 1.3×10^{-8} mutations per site per year for both markers (see Lessios, 2008) under a HKY model and a 0.25 inheritance scalar appropriate for mtDNA. An MCMC chain with a length of 1,000,000 sampled every 100 generations with 10% burn-in was used to estimate parameters for each species–gene combination. Five independent runs were computed to ensure convergence. The independent runs were subsampled and combined using the L mode of IMA2, and the median values of the parameter distributions for the combined runs are presented here. For *N. sammara* (COI) and *P. diacanthus* (COI and cyt *b*), which did not share any haplotypes between regions, prior values of *m* in both directions were set to zero. Although we regard all absolute parameter estimates with caution given that our data consist of two linked loci, and we apply mutation rates calibrated in other reef fishes, relative comparisons among species are likely to be robust to such approximations (Karl *et al.*, 2012).

RESULTS

Genetic diversity and population structure

259 COI and cyt *b* sequence data revealed divergent patterns of genetic diversity and
260 population structure among the seven sampled reef fish species. Haplotype (*h*)
261 and nucleotide (π) diversity was higher in four out of the seven species in the
262 WIO than in the Red Sea for COI, and five out of the seven species for cyt *b* when
263 populations within each region were pooled (Fig. 2, and see Appendix S1 in
264 Supporting Information). This trend cannot be explained by a greater sampling
265 effort in the WIO, given that species with comparable sample sizes for each
266 region, such as *Cephalopholis argus* and *P. diacanthus*, still had lower genetic
267 diversity in the Red Sea.

268 AMOVA supported the geographical grouping of sites into Red Sea and
269 WIO regions (Table 2). Although there was some variability in genetic
270 differentiation among sampling sites between regions (Appendix S2), six out of
271 the seven species showed significant genetic structure for at least one of the
272 molecular markers (Table 2).

273 Population pairwise tests were significantly different in 22 (for COI) or 16
274 (for cyt *b*) out of 47 comparisons (all $P \leq 0.01$); all significant comparisons were
275 between regions rather than between sites within regions, and ranged from 0.07 to
276 0.67 for Φ_{ST} and 0.05 to 1.00 for Jost's *D* (Fig. 3, Appendix S2). Estimates of
277 genetic differentiation across all species were correlated between molecular
278 markers (non-parametric Spearman's rank correlation: Φ_{ST}, $r = 0.77$, $P < 0.001$, *n*

279 = 47; Jost's *D*, $r = 0.57$, $P < 0.001$, $n = 47$), although the larger spread of Jost's *D*
280 values than Φ_{ST} values is probably related to the former not being constrained by
281 within-site heterozygosity. Pairwise genetic differentiation-based Φ_{ST} and Jost's
282 *D* were also significantly correlated across all datasets (non-parametric
283 Spearman's rank correlation coefficient: COI, $r = 0.67$, $P < 0.001$; cyt *b*, $r = 0.83$,
284 $P < 0.001$).

285 As expected from the Φ_{ST} values, the median-joining networks show more
286 shared COI or cyt *b* haplotypes between collection sites within the Red Sea and
287 WIO than between these regions (Fig. 4). The only exception to this pattern was
288 the high proportion of haplotypes shared between Al Lith or Thuwal (Red Sea)
289 and Oman (WIO) for *Cephalopholis argus.* Even though our comparisons
290 between Oman and other sites should be viewed with caution, given that these are
291 based on data from only a single species (*C. argus*) with a low sample size ($n = 8$
292 or 9), some endemic Red Sea fauna extend to the Omani coast (e.g. *Cirrhitus*
293 *spilotoceps*, M.R.G. & J.E. Randall, Bishop Museum, unpub. data).

294

295 **IMA2 analysis**

296 The estimated times since initial separation between the Red Sea and WIO
297 populations for the seven reef fish species ranged from approximately 21,000 to

830,000 years (Table 3). Recent separations of less than 100,000 years were apparent for *Chaetodon auriga* and *H. hortulanus*, older separations of 100,000 to 300,000 years were apparent for *A. nigrofuscus*, *Cephalopholis argus*, *L. kasmira*, and *N. sammara*, and finally *P. diacanthus* populations have been isolated for 660,000 to 830,000 years. Of the older separations, *L. kasmira* was characterized by high subsequent gene flow, whereas gene flow was restricted for *N. sammara* (and *P. diacanthus*); these two species also have the highest level of divergence between the Red Sea and WIO based on Φ_{ST} (Table 2). Differences among species in both the timing of initial divergence and subsequent migration rates reveal considerable variation in the link between Red Sea and WIO populations.

The direction of migration varied among species. For example, a higher proportion of migrants moved from the WIO into the Red Sea for *H. hortulanus*, whereas *L. kasmira* moved predominantly out of the Red Sea (Table 3). For the remaining species, gene flow was low in both directions, or driven by differences in effective population size, indicating no bias in effective migration between regions.

DISCUSSION

This study demonstrates barriers to gene flow between the Red Sea and WIO for some reef fish species, but an apparent lack of phylogeographical breaks for others, which may reflect the volatile geological history of the Red Sea region.

The Red Sea is a marginal water mass whose movement in the upper layers is driven by the summer and winter monsoons acting through a restricted connection with the adjacent Gulf of Aden (Siddall *et al.*, 2004; Biton *et al.*, 2008). During each glacial maximum of the Pleistocene, the last characterized by a 120-m drop in sea level only 20,000 years ago, the inflow pattern and exchange of surface water was limited owing to the shallow sill at the Straight of Bab al Mandab, the only natural gateway into the Red Sea (Siddall *et al.*, 2003). As a result, throughout these periods of isolation, increased evaporation may have raised temperature and salinity levels higher than most reef fishes can tolerate (e.g. > 50‰; Biton *et al.*, 2008), resulted in periods of reduced planktonic (i.e. larval) development (Hemleben *et al.*, 1996), and caused mass extirpation within the Red Sea (Sheppard *et al.*, 1992; but see Klausewitz, 1989).

In addition to intermittent historical barriers created by Pleistocene glacial cycles, contemporary barriers exist. The lack of coral habitat along the 2200 km coastline from Djibouti to southern Somalia may inhibit gene flow between the Red Sea and WIO by limiting opportunities for stepping stone dispersal (Kemp, 1998). Within the Red Sea, the extensive turbid-water region south of 20 °N may also inhibit larval dispersal or settlement (Ormond & Edwards, 1987; Roberts *et al.*, 1992). The long-term persistence and age of these contemporary barriers, however, is uncertain.

339 Most genetic work on reef fishes within the Red Sea has focused on the
340 differentiation of fauna between the Gulf of Aqaba and northern Red Sea, with
341 some notable exceptions. Froukh & Kochzius (2008) identified a damselfish in
342 the Red Sea (*Chromis viridis*) as being distinct from conspecifics in Indonesia and
343 the Philippines based on mtDNA sequences. Similar research on marine
344 invertebrates (*Acanthaster planci*: Benzie *et al.*, 1999; *Scylla serrata*: Fratini &
345 Vannini, 2002) support a genetic distinction of the Red Sea populations. In
346 contrast, Kochzius & Blohm (2005) found no mtDNA differentation between
347 lionfish (*Pterois miles*) populations in the Red Sea and Indian Ocean.
348 Five of the seven species we examined were genetically differentiated
349 between the Red Sea and WIO based on AMOVA and median-joining networks.
350 *Halichoeres hortulanus* and *L. kasmira* had minimal or inconsistent genetic
351 differentiation, as well as extensive mixing of haplotypes within and between
352 regions. *Acanthurus nigrofuscus, Cephalopholis argus,* and *Chaetodon auriga*
353 had modest differentiation between regions with pronounced separation of
354 peripheral haplotypes, but shared a common haplotype among all sampling sites.
355 *Neoniphon sammara* and *P. diacanthus* had fixed differences between regions.
356 Variability in genetic signatures can occur even among closely related
357 species (Rocha *et al.*, 2002; Gaither *et al.*, 2010; DiBattista *et al.*, 2012b) and may
358 be related to innate differences in life-history or ecological preferences, although
359 these widely distributed species are all presumably capable of long distance

dispersal (e.g. Eble *et al.*, 2009, 2011; Gaither *et al.*, 2010, 2011) based on available estimates of pelagic larval duration (range: 24.5–48 days; Thresher & Brothers, 1985; Victor, 1986; Wilson & McCormick, 1999) and longitudinal range size (range: 20,063–21,689 km; Randall, 1999, 2005). Indeed, our study species cover a wide spectrum of dietary modes ranging from herbivory (*A. nigrofuscus*) to specialist feeding on sessile or mobile invertebrates (*Chaetodon auriga*, *H. hortulanus*, and *P. diacanthus*) to piscivory (*Cephalopholis argus*, *L. kasmira*, and *N. sammara*). These species also display a variety of reproductive behaviour ranging from dioecism (*L. kasmira*, *N. sammara*) with mate-pairing (*Chaetodon auriga*) or spawning aggregations (*A. nigrofuscus*) to protogyny (*Cephalopholis argus* and *H. hortulanus*). Given that there are no real unifying life-history features for this diverse group, we suspect that differences in ecological resilience to geological disturbance may have contributed to the range of colonization histories, although this will require further testing.

Considering the prevailing currents in the Indian Ocean, it is not surprising that sampling sites in the WIO were genetically similar to each other. The Chagos and Seychelles archipelagos are located in the South Equatorial Current, which flows from east to west. Both archipelagos are also heavily influenced by seasonal or permanent countercurrents (South Equatorial Countercurrent and East African Countercurrent, respectively; Fig. 1). The strong and variable water movement of the region has resulted in Diego Garcia, which is located at the southern end of

the Chagos Archipelago, having faunal affinities with both the Indo-Polynesian and WIO provinces (Winterbottom & Anderson, 1997; Craig, 2008; Gaither *et al.*, 2011; Briggs & Bowen, 2012). The South African coastline is similarly well connected to the central Indian Ocean, being influenced by the warm Mozambique/Agulhas current (Lutjeharms, 2006), which facilitates unidirectional (north to south) transport of tropical fauna from other sites in the WIO.

There are several records of long distance dispersal of tropical reef fish (e.g. *Chaetodon zanzibarensis* and *Ecsenius lineatus*) to the Arabian coastline during periods of upwelling, which indicate that larval transport from the WIO to this region and subsequent settlement are not precluded (Kemp, 2000). Although we only sampled a few specimens (n = 9) of a single species off the coast of Oman (*Cephalopholis argus*), these fish were not genetically distinct from conspecifics sampled at Diego Garcia (COI: Φ_{ST} = 0.029, P = 0.25; cyt *b*: Φ_{ST} < 0.001, P = 0.50) or the Seychelles (COI: Φ_{ST} <0.001, P = 0.80; cyt *b*: Φ_{ST} = 0.039, P = 0.20).

Vicariance events and colonization history

Our mtDNA data provide evidence for three separate periods of colonization or export of propagules between the Red Sea and WIO (Table 3). First, Red Sea populations of *Chaetodon auriga*, and *H. hortulanus* appear to derive from the WIO during or soon after the most recent glacial maximum (c.21,000 to 31,000

years ago; but see Karl *et al.*, 2012). Second, population separations in *A. nigrofuscus*, *Cephalopholis argus*, and *L. kasmira* pre-date the Last Glacial Maximum but include recurrent gene flow in most cases. Third, *N. sammara* and *P. diacanthus* represent long-isolated evolutionary lineages in the Red Sea. These final cases in particular indicate that some Red Sea residents survived the major temperature and salinity crises recorded 19,000, 30,000 and 450,000 years ago (Siddall *et al.*, 2003).

IMA2 analyses indicate bidirectional gene flow between the Red Sea and WIO, which is also apparent in the older history inscribed in haplotype networks (Fig. 4). Three cases provide especially strong inference: (1) in the COI and cyt *b* network for *A. nigrofuscus*, the central (ancestral) haplotype is observed primarily in the Indian Ocean, whereas the Red Sea haplotypes are peripheral; (2) in the COI network for *N. sammara*, the central haplotype is detected only in the Red Sea, with the Indian Ocean haplotypes peripheral; and (3) in the cyt *b* network for *P. diacanthus*, the central haplotype is detected only in the Red Sea. Hence the networks for these three species indicate colonization into and out of the Red Sea, which supports the hypothesis that peripheral habitats can export biodiversity to the central Indo-Pacific.

Taxonomic considerations

Our genetic study highlights three interesting cases where the current classification of existing species may not reflect their evolutionary history. *Chaetodon auriga* is one of the most widespread butterflyfishes on the planet, with a distribution of approximately 82.2 million km^2 across the tropical Indo-Pacific (Allen *et al.*, 1998). The original species description is from Red Sea specimens, which lack a dark spot on the margin of the soft dorsal fin, such that conspecifics outside the Red Sea were regarded as the subspecies *Chaetodon auriga setifer* (Bloch, 1795).

Although we did detect differences in mtDNA sequences between the Red Sea and WIO, these were only marginally significant. In addition, the most common haplotype was shared between the Red Sea and WIO at both COI and cyt *b*. Even though colour morphs do correspond to genetic partitions in some species (e.g. Craig & Randall, 2008; Drew *et al.*, 2008; Randall & Rocha, 2009), discordance between genetic divergence and coloration is well documented in reef fishes (Ramon *et al.*, 2003; Rocha, 2004; Messmer *et al.*, 2005), including butterflyfishes (family Chaetodontidae: McMillan & Palumbi, 1995) and their sister group angelfishes (family Pomacanthidae: Bowen *et al.*, 2006; DiBattista *et al.*, 2012a). For these reasons, we regard the Red Sea population as conspecific with all other populations of *Chaetodon auriga*, although the shallow but significant population genetic differentiation supports the sub-specific status.

Two species in our study reveal a strikingly different pattern: *N. sammara* and *P. diacanthus* were characterized by high Φ_{ST} (or Jost's *D*) values relative to all other species and strong mtDNA differences between regions. For *P. diacanthus* in particular, Red Sea and WIO haplotypes are separated by at least three mutations at cyt *b* and one fixed mutation at COI, indicating isolation for several hundred thousand years. This genetic separation is matched by coloration differences between Red Sea and WIO populations (Allen *et al.*, 1998), indicating long-isolated populations that, unlike other species examined, have failed to reconnect during interglacial periods. Notably the species is absent from sites in the Arabian Sea, indicating geographic isolation (Kemp, 1998). While we know of no coloration or morphological differences in *N. sammara* that may indicate cryptic lineages, this possibility merits further investigation.

CONCLUSIONS

Comparative phylogeographical studies have done much to illuminate the evolutionary history of regional marine faunas (Avise, 1992; Lessios & Robertson, 2006; Kelly & Palumbi, 2010; Carpenter *et al.*, 2011; Toonen *et al.*, 2011; Ludt *et al.*, 2012). Here we provide the first multispecies comparison between Red Sea and Indian Ocean reef fishes, and find a spectrum of outcomes from recent gene flow to ancient evolutionary separations. Three broad conclusions are apparent. First, endemism and biodiversity are higher among Red

463 Sea reef fishes than previously suspected (i.e. *N. sammara* and *P. diacanthus*),
464 and ongoing studies will be likely to elevate these estimates. Second, some
465 elements of the Red Sea fauna survived the salinity crises caused by late
466 Pleistocene glaciations. This does not require continuous residence in the Red
467 Sea, as persistence in the Gulf of Aden just outside the Red Sea remains a
468 possibility. It seems unlikely, however, that a genetically distinct and cohesive
469 fauna could survive in the Gulf without extensive admixture with other Indian
470 Ocean populations. We therefore favour the explanation that Red Sea refugia
471 existed during low sea level stands associated with glaciations. Third, peripheral
472 habitats such as marginal seas and isolated archipelagos are not necessarily
473 'evolutionary graveyards'. Rather, our data indicate that the Red Sea is capable of
474 exporting biodiversity to the broader Indo-Pacific, thus operating as a potential
475 engine of evolutionary diversity in our oceans.

476

477

478

479

480

ACKNOWLEDGEMENTS

This research was supported by National Science Foundation grants OCE-0453167 and OCE-0929031 to B.W.B., National Geographic Society Grant 9024-11 to J.D.D., KAUST Red Sea Research Center funding to M.L.B., California Academy of Sciences funding to L.A.R., and by a Natural Sciences and Engineering Research Council of Canada (NSERC) postgraduate fellowship to J.D.D. For specimen collections, we thank Gavin Gouws (South Africa Institute for Aquatic Biodiversity), Matthew Iacchei, Kelton W. McMahon, Gerrit Nanninga, Jonathan Puritz, and Charles R.C. Sheppard. We also thank Robert J. Toonen, Serge Planes, John E. Randall, Claudia Rocha, Jo-Ann C. Leong, Eric Mason at Dream Divers, David Pence, the KAUST Coastal and Marine Resources Core Lab, the Administration of the British Indian Ocean Territory, and members of the ToBo lab for logistic support; we thank Stephan Moldzio for photos of *Neoniphon sammara*; we thank the Center for Genomics, Proteomics, and Bioinformatics at the University of Hawai‘i at Mānoa, in addition to the KAUST Bioscience Core Facility for their assistance with DNA sequencing. This is contribution no. 1530 from the Hawai‘i Institute of Marine Biology and no. 8790 from the School of Ocean and Earth Science and Technology.

REFERENCES

Allen, G.R., Steene, R. & Allen, M. (1998) *A guide to angelfishes and butterflyfishes*. Odyssey Publishing, Perth.

Avise, J.C. (1992) Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos*, **3,** 62–76.

Bailey, G. (2009) The Red Sea, coastal landscapes, and hominin dispersals. *The evolution of human populations in Arabia* (ed. by M. Petraglia and J. Rose), pp. 15–37. Springer, Dordrecht.

Bandelt, H.J., Forster, P. & Röhl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16,** 37–48.

Benzie, J.A.H. (1999) Major genetic differences between crown-of-thorns starfish (*Acanthaster planci*) populations in the Indian and Pacific Oceans. *Evolution*, **53,** 1782–1795.

Biton, E., Gildor, H. & Peltier, W.R. (2008) Red Sea during the Last Glacial Maximum: implications for sea level reconstruction. *Paleoceanography*, **23,** PA1214.

Biton, E., Gildor, H., Trommer, G., Siccha, M., Kucera, M., van der Meer, M.T.J. & Schouten, S. (2010) Sensitivity of Red Sea circulation to monsoonal

variability during the Holocene: an integrated data and modeling study. *Paleoceanography*, **25,** PA4209.

Bosworth, W., Huchon, P. & McClay, K. (2005) The Red Sea and Gulf of Aden Basins. *Journal of African Earth Sciences*, **43,** 334–378.

Bowen, B.W., Bass, A.L., Rocha, L.A., Grant, W.S. & Robertson, D.R. (2001) Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global scale. *Evolution*, **55,** 1029–1039.

Bowen, B.W., Muss, A., Rocha, L.A. & Grant W.S. (2006) Shallow mtDNA coalescence in Atlantic pygmy angelfishes (genus *Centropyge*) indicates a recent invasion from the Indian Ocean. *Journal of Heredity*, **97,** 1–12.

Briggs, J.C. (1999) Coincident biogeographic patterns: Indo-West Pacific Ocean. *Evolution*, **53,** 326–335.

Briggs, J.C. & Bowen, B.W. (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography*, **39,** 12–30.

Carpenter, K.E., Barber, P.H., Crandall, E.D., Ablan-Lagman, M.C.A., Ambariyanto, Mahardika, G.N., Manjaji-Matsumoto, B.M., Juinio-Meñez, M.A., Santos, M.D., Starger, C.J. & Toha, A.H.A. (2011) Comparative phylogeography of the Coral Triangle and implications for marine management. *Journal of Marine Biology*, **2011,** 396982.

Chao, A., Jost, L., Chiang, S.C., Jiang, Y-H. & Chazdon, R.L. (2008) A two stage probabilistic approach to multiple-community similarity indices. *Biometrics*, **64,** 1178–1186.

Cox, C.B. & Moore, P.D. (2000) *Biogeography: an ecological and evolutionary approach*, 6th edn. Blackwell, Oxford.

Craig, M.T. (2008) The goldrim surgeonfish (*Acanthurus nigricans*; Acanthuridae) from Diego Garcia, Chagos Archipelago: first record for the central Indian Ocean. *Zootaxa*, **1850,** 65–68.

Craig, M.T. & Randall, J.E. (2008) Two new species of the Indo-Pacific clingfish genus *Discotrema* (Gobiesocidae). *Copeia*, **2008,** 68–74.

DeVantier, L.M., Turak, E., Al-Shaikh, K.A. & De'ath, G. (2000) Coral communities of the central-northern Saudi Arabian Red Sea. *Fauna of Arabia*, **18,** 23–66.

DiBattista, J.D., Wilcox, C., Craig, M.T., Rocha, L.A. & Bowen, B.W. (2011) Phylogeography of the Pacific Blueline Surgeonfish, *Acanthurus nigrori,s*, reveals a cryptic species in the Hawaiian archipelago. Journal of Marine Biology Article ID 839134.

DiBattista, J.D., Waldrop, E., Bowen, B.W., Schultz, J.K., Gaither, M.R., Pyle, R.L. & Rocha, L.A. (2012a) Twisted sister species of pygmy angelfishes: discordance between taxonomy, coloration, and phylogenetics. *Coral Reefs*, **31,** 839–851.

561 DiBattista, J.D., Rocha, L.A., Craig, M.T., Feldheim, K.A. & Bowen, B.W.
562 (2012b) Phylogeography of two closely related Indo-Pacific
563 butterflyfishes reveals divergent evolutionary histories and discordant
564 results from mtDNA and microsatellites. *Journal of Heredity*, **103,** 617–
565 629.
566 Drew, J., Allen, G.R., Kaufman, L. & Barber, P.H. (2008) Endemism and regional
567 color and genetic differences in five putatively cosmopolitan reef fishes.
568 *Conservation Biology*, **22,** 965–975.

569 Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R.,
570 Stones-Havas, S., Thierer, T. & Wilson, A. (2009) *Geneious v4.8.*
571 Available at: http://www.geneious.com/.

572 Eble, J.A., Toonen, R.J. & Bowen, B.W. (2009) Endemism and dispersal:
573 comparative phylogeography of three surgeonfishes across the Hawaiian
574 Archipelago. *Marine Biology*, **156,** 689–698.
575 Eble, J.A., Toonen, R.J., Sorenson, L., Basch, L.V., Papastamatiou, Y.P. &
576 Bowen, B.W. (2011) Escaping paradise: larval export from Hawaii in an
577 Indo-Pacific reef fish, the yellow tang (*Zebrasoma flavescens*). *Marine*
578 *Ecology Progress Series*, **428,** 245–258.
579 Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular
580 variance inferred from metric distances among DNA haplotypes:

application to human mitochondrial DNA restriction data. *Genetics*, **131,** 479–491.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1,** 47–50.

Forsskål, P. (1775) Descriptiones animalium avium, amphibiorum, piscium, insectorum, vermium; quae in itinere orientali observavit. Möller, Copenhagen, Denmark.

Fratini, S. & Vannini, M. (2002) Genetic differentiation in the mud crab *Scylla serrata* (Decapoda: Portunidae) within the Indian Ocean. *Journal of Experimental Marine Biology and Ecology*, **272,** 103–116.

Froukh, T. & Kochzius, M. (2007) Genetic population structure of the endemic fourline wrasse (*Larabicus quadrilineatus*) suggests limited larval dispersal distances in the Red Sea. *Molecular Ecology*, **16,** 1359–1367.

Froukh, T. & Kochzius, M. (2008) Species boundaries and evolutionary lineages in the blue green damselfishes *Chromis viridis* and *Chromis atripectoralis* (Pomacentridae). *Journal of Fish Biology*, **72,** 451–457.

Gaither, M.R., Toonen, R.J., Robertson, D.R., Planes, S. & Bowen, B.W. (2010) Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). *Journal of Biogeography*, **37**, 133–147.

602 Gaither, M.R., Bowen, B.W., Bordenave, T.-R., Rocha, L.A., Newman, S.J.,
603 Gomez, J.A., van Herwerden, L. & Craig, M.T. (2011) Phylogeography of
604 the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene
605 isolation across the Indo-Pacific barrier with contemporary overlap in the
606 Coral Triangle. *BMC Evolutionary Biology*, **11,** 189.
607 Golani, D. & Bogorodsky, S.V. (2010) The fishes of the Red Sea—reappraisal
608 and updated checklist. *Zootaxa*, **2463,** 1–135.
609 Hasegawa, M., Kishino, H. & Yano, T-A. (1985) Dating of the human-ape
610 splitting by a molecular clock of mitochondrial DNA. *Journal of*
611 *Molecular Evolution*, **22,** 160–174.
612 Hassan, M., Harmelin-Vivien, M. & Bonhomme, F. (2003) Lessepsian invasion
613 without bottleneck: example of two rabbitfish species (*Siganus rivulatus*
614 and *Siganus luridus*). *Journal of Experimental Marine Biology and*
615 *Ecology*, **291,** 219–232.
616 Hemleben, C., Meishchner, D., Zahn, R., Almogi-Labin, A., Erlenkeuser, H. &
617 Hiller, B. (1996) Three hundred eighty thousand year long stable isotope
618 and faunal records from the Red Sea: influence of global sea level change
619 on hydrography. *Paleoceanography*, **11,** 147–156.
620 Hey, J. (2010) Isolation with migration models for more than two populations.
621 *Molecular Biology and Evolution*, **27,** 905–920.

622 Hey, J. & Nielsen, R. (2007) Integration within the Felsenstein equation for
623 improved Markov chain Monte Carlo methods in population genetics.
624 *Proceedings of the National Academy of Sciences USA*, **104**, 2785–2790.
625 Jost, L. (2008) G_{ST} and its relatives do not measure differentiation. *Molecular*
626 *Ecology*, **17,** 4015–4026.
627 Karl, S.A., Toonen, R.J., Grant, W.S. & Bowen, B.W. (2012) Common
628 misconceptions in molecular ecology: echoes of the modern synthesis.
629 *Molecular Ecology*, **21,** 4171–4189.
630 Kelly, R.P. & Palumbi, S.R. (2010) Genetic structure among 50 species of the
631 northeastern Pacific rocky intertidal community. *PLoS ONE*, **5,** e8594.
632 Kemp, J. (1998) Zoogeography of the coral reef fishes of the Socotra
633 Archipelago. *Journal of Biogeography*, **25,** 919–933.
634 Kemp, J. (2000) Zoogeography of the coral reef fishes of the northeastern Gulf of
635 Aden, with eight new records of coral reef fishes from Arabia. *Fauna of*
636 *Arabia*, **18,** 293–321.
637 Klausewitz, W. (1989) Evolutionary history and zoogeography of the Red Sea
638 ichthyofauna. *Fauna of Saudi Arabia*, **10**, 310–337.
639 Kochzius, M. & Blohm, D. (2005) Genetic population structure of the lionfish
640 *Pterois miles* (Scorpaenidae, Pteroinae) in the Gulf of Aqaba and northern
641 Red Sea. *Gene*, **347,** 295–301.

642 Lessios, H.A. (2008) The Great American Schism: divergence of marine
643 organisms after the rise of the Central American isthmus. *Annual Review*
644 *of Ecology, Evolution, and Systematics*, **39,** 63–91.
645 Lessios, H.A. & Robertson, D.R. (2006) Crossing the impassible: genetic
646 connections in 20 reef fishes across the eastern Pacific barrier.
647 *Proceedings of the Royal Society B: Biological Sciences*, **273,** 2201–2208.
648 Ludt, W.B., Bernal, M., Bowen, B.W. & Rocha L.A. (2012) Living in the past:
649 phylogeography and population histories of Indo-Pacific wrasses (genus
650 *Halichoeres*) in shallow lagoons versus outer reef slopes. *PLoS ONE*, **7,**
651 e38042.
652 Lutjeharms, J.R.E. (2006) *The Agulhas Current*. Springer, New York, NY.
653 McMillan, W.O. & Palumbi, S.R. (1995) Concordant evolutionary patterns
654 among Indo-West Pacific butterflyfishes. *Proceedings of the Royal Society*
655 *B: Biological Sciences*, **260,** 229–236.
656 Meeker, N.D., Hutchinson, S.A., Ho, L. & Trede, N.S. (2007) Method for
657 isolation of PCR-ready genomic DNA from zebrafish tissues.
658 *Biotechniques*, **43,** 610–614.
659 Messmer, V., van Herwerden, L., Munday, P.L. & Jones, G.P. (2005)
660 Phylogeography of colour polymorphism in the coral reef fish
661 *Pseudochromis fuscus*, from Papua New Guinea and the Great Barrier
662 Reef. *Coral Reefs*, **24,** 392–402.

663 Meyer, A. (1994) Shortcomings of the cytochrome *b* gene as a molecular marker.
664 *Trends in Ecology and Evolution*, **9,** 278–280.
665 Narum, S.R. (2006) Beyond Bonferroni: less conservative analyses for
666 conservation genetics. *Conservation Genetics*, **7,** 783–787.
667 Ngugi, D.K., Antunes, A., Brune, A. & Stingl, U. (2012) Biogeography of pelagic
668 bacterioplankton across an antagonistic temperature–salinity gradient in
669 the Red Sea. *Molecular Ecology*, **21,** 388–405.
670 Ormond, R. & Edwards, A. (1987) Red Sea fishes. *Key environments – Red Sea*
671 (ed. by A.J. Edwards and S.M. Head), pp. 251–287. Pergamon Press,
672 Oxford.
673 Posada, D. (2008) jModelTest: phylogenetic model averaging. *Molecular Biology*
674 *and Evolution*, **25,** 1253–1256.
675 Raitsos, D.E., Hoteit, I., Prihartato, P.K., Chronis, T., Triantafyllou G. &
676 Abualnaja, Y. (2011) Abrupt warming of the Red Sea. *Geophysical*
677 *Research Letters*, **38,** L14601.
678 Ramon, M.L., Lobel, P.S. & Sorenson, M.D. (2003) Lack of mitochondrial
679 genetic structure in hamlets (*Hypoplectrus* spp.): recent speciation or
680 ongoing hybridization? *Molecular Ecology*, **12,** 2975–2980.
681 Randall, J.E. (1994) Twenty-two new records of fishes from the Red Sea. *Fauna*
682 *of Saudi Arabia*, **14,** 259–275.

683 Randall, J.E. (1999) Report on fish collections from the Pitcairn Islands. *Atoll*
684 *Research Bulletin*, **461**, 1–53.

685 Randall, J.E. (2005) *Reef and shore fishes of the South Pacific: New Caledonia to*
686 *Tahiti and the Pitcairn Islands*. University of Hawai'i Press, Honolulu, HI.

687 Randall, J.E. & Rocha, L.A. (2009) *Chaetodontoplus poliourus*, a new angelfish
688 (Perciformes: Pomacanthidae) from the tropical western Pacific. *The*
689 *Raffles Bulletin of Zoology*, **57,** 511–520.
690 Reece, J.S., Bowen, B.W., Smith, D.G. & Larson A. (2010) Molecular
691 phylogenetics of moray eels (Muraenidae) demonstrates multiple origins
692 of a shell-crushing jaw (*Gymnomuraena*, *Echidna*) and multiple
693 colonizations of the Atlantic Ocean. *Molecular Phylogenetics and*
694 *Evolution*, **57,** 829–835.
695 Riegl, B. & Velimirov, B. (1994) The structure of coral communities at Hurghada
696 in the Northern Red Sea. *Marine Ecology*, **15,** 213–231.
697 Righton, D., Kemp, J. & Ormond, R. (1996) Biogeography, community structure
698 and diversity of Red Sea and western Indian Ocean butterflyfishes.
699 *Journal of the Marine Biological Association of the UK*, **76,** 223–228.
700 Roberts, C.M., Shepherd, A.R.D. & Ormond, R.F.G. (1992) Large-scale variation
701 in assemblage structure of Red Sea butterflyfishes and angelfishes.
702 *Journal of Biogeography*, **19,** 239–250.

703 Rocha, L.A. (2004) Mitochondrial DNA and color pattern variation in three
704 western Atlantic *Halichoeres* (Labridae), with the revalidation of two
705 species. *Copeia*, **2004,** 770–782.
706 Rocha, L.A., Bass, A.L., Robertson, D.R. & Bowen, B.W. (2002) Adult habitat
707 preferences, larval dispersal, and the comparative phylogeography of three
708 Atlantic surgeonfishes (Teleostei: Acanthuridae). *Molecular Ecology*, **11,**
709 243–251.
710 Seutin, G., White, B.N. & Boag, P.T. (1991) Preservation of avian blood and
711 tissue samples for DNA analyses. *Canadian Journal of Zoology*, **69,** 82–
712 90.
713 Sheppard, C.R.C. (2000) Coral reefs of the western Indian Ocean: an overview.
714 *Coral reefs of the Indian Ocean: their ecology and conservation* (ed. by
715 T.R. McClanahan, C.R.C. Sheppard and D.O. Obura), pp. 3–38. Oxford
716 University Press, Oxford.
717 Sheppard, C.R.C. & Sheppard, A.L.S. (1991) Corals and coral communities of
718 Arabia. *Fauna of Saudi Arabia*, **12**, 3–170.
719 Sheppard, C.R.C., Price, A.R.G. & Roberts, C.M. (1992) *Marine ecology of the*
720 *Arabian region: patterns and processes in extreme tropical environments*.
721 Academic Press, London.

722 Siddall, M., Rohling, E.J., Almogi-Labin, A., Hemleben, C., Meischner, D.,
723 Schmelzer, I. & Smeed, D.A. (2003) Sea-level fluctuations during the last
724 glacial cycle. *Nature*, **423,** 853–858.
725 Siddall, M., Smeed, D.A., Hemleben, C., Rohling, E.J., Schmelzer, I. & Peltier,
726 W.R. (2004) Understanding the Red Sea response to sea level. *Earth and*
727 *Planetary Science Letters*, **225,** 421– 434.
728 Skillings, D.J., Bird, C.E., & Toonen, R.J. (2011) Gateways to Hawai'i: genetic
729 population structure of the tropical sea cucumber *Holothuria atra*. *Journal*
730 *of Marine Biology*, **2011,** 783030.
731 Smeed, D. (2004) Exchange through the Bab el Mandab. *Deep-Sea Research Part*
732 *II: Tropical Studies in Oceanography*, **51,** 455–474.
733 Smeed, D.A. (1997) Seasonal variation of the flow in the strait of Bab al Mandab.
734 *Oceanologica Acta*, **20,** 773–781.
735 Song, C.B., Near, T.J. & Page, L.M. (1998) Phylogenetic relations among percid
736 fishes as inferred from mitochondrial cytochrome *b* DNA sequence data.
737 *Molecular Phylogenetics and Evolution*, **10,** 343–353.
738 Taberlet, P., Meyer, A. & Bouvet, J. (1992) Unusual mitochondrial variation in
739 two local populations of blue tit *Parus caeruleus*. *Molecular Ecology*, **1,**
740 27–36.

741 Thresher, R.E. & Brothers, E.B. (1985) Reproductive ecology and biogeography
742 of Indo-West Pacific angelfish (Pisces: Pomacanthidae). *Evolution,* **39,**
743 878–887.
744 Toonen, R.J., Andrews, K.R., Baums, I.B., Bird, C.E., Concepcion, G.T., Daly-
745 Engel, T.S., Eble, J.A., Faucci, A., Gaither, M.R., Iacchei, M., Puritz, J.B.,
746 Schultz, J.K., Skillings, D.J., Timmers, M.A. & Bowen, B.W. (2011)
747 Defining boundaries for ecosystem-based management: a multispecies
748 case study of marine connectivity across the Hawaiian archipelago.
749 *Journal of Marine Biology*, **2011,** 460173.
750 Victor, B.C. (1986) Duration of the planktonic larval stage of one hundred species
751 of Pacific and Atlantic wrasses (family Labridae). *Marine Biology*, **90,**
752 317–326.
753 Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R. & Hebert, P.D.N. (2005) DNA
754 barcoding Australia's fish species. *Philosophical Transactions of the*
755 *Royal Society B: Biological Sciences*, **360,** 1847–1857.
756 Wilson, D.T. & McCormick, M.I. (1999) Microstructure of settlement-marks in
757 the otoliths of tropical reef fish. *Marine Biology*, **134,** 29–41.

758 Winterbottom, R. & Anderson, R.C. (1997) A revised checklist of the epipelagic
759 and shore fishes of the Chagos Archipelago, Central Indian Ocean. *Ichthyological*
760 *Bulletin*, **66,** 1–28.

761 **SUPPORTING INFORMATION**

762 Additional Supporting Information may be found in the online version of this
763 article:

764 **Appendix S1** Mitochondrial DNA haplotype (*h*) and nucleotide diversity (π) for
765 each species from each collection location.

766 **Appendix S2** Population pairwise Φ_{ST} and Jost's *D* values for each species based
767 on mitochondrial DNA sequences.

768 As a service to our authors and readers, this journal provides supporting
769 information supplied by the authors. Such materials are peer-reviewed and may be
770 reorganized for online delivery, but are not copy-edited or typeset. Technical
771 support issues arising from supporting information (other than missing files)
772 should be addressed to the authors.

773

774

775

776

Biosketch

The authors' interests are focused on illuminating the evolutionary processes that generate marine biodiversity. They have carried out phylogeographical surveys of over 20 reef fish species in the greater Indo-Pacific to test existing evolutionary models, resolve the life history traits that influence dispersal and population separations in reef organisms, and inform marine conservation (e.g. defining the boundaries of marine protected areas).

Author contributions: J.D.D. conceived the ideas for this study, collected tissue samples and produced DNA sequences, analysed the data, and led the writing. In addition to contributing to writing, M.L.B., B.W.B., J.H.C. and M.T.C. collected tissue samples, M.R.G. and L.A.R. collected tissue samples and produced DNA sequences, J.A.E. produced DNA sequences, and D.J.S. implemented and interpreted coalescent analyses.

Editor: Craig McClain

Table 1 Study species, number of specimens, fragment length, primers used, and annealing temperatures for mitochondrial DNA cytochrome-c oxidase subunit-I (COI) and cytochrome *b* (cyt *b*) genes. DNA sequences from each primary collection location (Al Lith, Thuwal, Diego Garcia, and the Republic of Seychelles; see text) are described, along with collections made opportunistically at additional locations in the Western Indian Ocean (WIO). All haplotypes are available online in GenBank (accession numbers: KC187734-KC188056).

Species	Molecular sequence data								
			Red Sea		WIO				
	DNA fragment	Fragment length (bp)	Al Lith	Thuwal	Diego Garcia	Seychelles	Other sites (*n*)	Primer set	Annealing temp. (°C)
Acanthurus nigrofuscus	COI	634	22	27	31	31	–	Fish F2–Fish R2 (1)	50
(brown surgeonfish)	cyt *b*	683	22	28	31	30	–	Cyb9-–Cyb7 (2,3)	58
Cephalopholis argus	COI	537	26	19	24	10	Oman (8)	Fish F2–Fish R2 (1)	56
(peacock hind)	cyt *b*	618	27	22	32	13	Oman (9)	CB6F–CB6R (4)	54
Chaetodon auriga	COI	625	27	20	33	30	–	Fish F2–Fish R2 (1)	52
(threadfin butterflyfish)	cyt *b*	670	27	20	33	30	–	Cyb9–Cyb7 (2,3)	56
Halichoeres hortulanus	COI	589	25	27	20	28	–	Fish F2–Fish R2 (1)	50
(checkerboard wrasse)	cyt *b*	692	25	27	27	22	–	Cyb9–Cyb7 (2,3)	50
Lutjanus kasmira	COI	606	21	22	33	20	Sodwana Bay (34)	Fish F2–Fish R2 (1)	56
(bluestripe snapper)	cyt *b*	475	23	23	34	19	Sodwana Bay (34)	H15020–Cyb5 (5,3)	48
Neoniphon sammara	COI	611	20	31	30	28	–	Fish F2–Fish R2 (1)	50
(Sammara squirrelfish)	cyt *b*	508	20	31	29	38	–	NSAFOR4–NSAREV4*	60
Pygoplites diacanthus	COI	634	24	23	33	–	–	Fish F2–Fish R2 (1)	50
(regal angelfish)	cyt *b*	640	24	23	32	–	–	PydCytbF3–PydCytbR4*	50

*We designed two novel primer sets to amplify and sequence cyt *b* for *Neoniphon sammara* and *Pygoplites diacanthus*. Their sequences were as follows: NSAFOR4: 5′–TGC CGT GAC GTA AAC TAT GG–3′; NSAREV4: 5′–TGA AGT TGT CGG GAT CTC CT–3′; PydCytbF3: 5′–ATG GCA AAC TTA CGC AAA ACC–3′; PydCytbR4: 5′–GGC TGG TGT GAA GTT GTC–3′.

References: (1) Ward *et al.*, 2005; (2) Song *et al.*, 1998; (3) Taberlet *et al.*, 1992; (4) Gaither *et al.*, 2010; (5) Meyer, 1994.

Table 2 Analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) comparing variation between the Red Sea and Western Indian Ocean populations of reef fish based on mitochondrial DNA cytochrome-c oxidase subunit-I (COI) and cytochrome *b* (cyt *b*) genes. Site-specific samples sizes are shown in Table 1.

Species	DNA fragment	Percentage variation			Overall Φ_{ST}	*P*	Jost's *D* (SE)
		within populations	between populations	between regions			
Acanthurus nigrofuscus (brown surgeonfish)	COI	76.58	0.25	23.18	0.23	< 0.001	0.59 (0.051)
	cyt *b*	87.96	−0.64	12.68	0.12	< 0.001	0.21 (0.087)
Cephalopholis argus (peacock hind)	COI	72.45	−0.31	27.86	0.28	< 0.001	0.13 (0.088)
	cyt *b*	77.59	1.02	21.39	0.22	< 0.001	0.13 (0.060)
Chaetodon auriga (threadfin butterflyfish)	COI	80.35	−1.80	21.45	0.20	< 0.001	0.087 (0.041)
	cyt *b*	98.48	−0.44	1.96	0.015	0.12	0.021 (0.022)
Halichoeres hortulanus (checkerboard wrasse)	COI	97.30	2.60	0.10	0.027	0.071	0.041 (0.019)
	cyt *b*	95.44	1.08	3.48	0.046	0.045	0.078 (0.088)
Lutjanus kasmira (bluestripe snapper)	COI	100.11	−1.30	1.20	0	0.83	0.006 (0.023)
	cyt *b*	98.88	−0.46	1.58	0.011	0.23	0.090 (0.059)
Neoniphon sammara (Sammara squirrelfish)	COI	66.03	−0.33	34.30	0.34	< 0.001	0.68 (0.038)
	cyt *b*	70.28	−0.69	30.41	0.30	< 0.001	0.59 (0.060)
Pygoplites diacanthus (regal angelfish)	COI	30.61	−1.17	70.57	0.69	< 0.001	0.61 (0.015)
	cyt *b*	65.00	−2.30	37.30	0.35	< 0.001	0.60 (0.027)

Table 3 Estimates of time in years (*t*) since initial separation, effective migration rate ($2N_em$), effective population sizes (N_e), and mutation-scaled migrations rates (*m*) between Red Sea (RS) and Western Indian Ocean (WIO) populations of seven reef fish species based on mitochondrial DNA cytochrome-c oxidase subunit-I (COI) and cytochrome *b* (cyt *b*) runs in IMA2 (Hey & Nielsen, 2007). Abbreviations: *NC*, no convergence. Inequalities: posterior probability densities rise to a plateau, so that all estimates larger than the shown value have the same approximate posterior probability.

Species	DNA fragment	Initial separation in years (*t*)	Effective migration rate ($2N_em$)		Effective population size (N_e)		Migration (*m*)	
			WIO to RS	RS to WIO	WIO	RS	WIO to RS	RS to WIO
Acanthurus nigrofuscus	COI	105,000	1.21	0.09	25.10	50.70	0.03	0.01
(brown surgeonfish)	cyt *b*	79,000	3.92	0.11	52.75	85.25	0.02	0.001
Cephalopholis argus	COI	> 121,000	NC	0.73	0.25	NC	21.05	1.45
(peacock bind)	cyt *b*	> 212,000	0.01	NC	NC	0.60	0.01	6.27
Chaetodon auriga	COI	26,800	0.08	0.13	6.60	3.80	0.01	0.01
(threadfin butterflyfish)	cyt *b*	30,700	0.03	4.85	34.65	1.65	0.01	0.07
Halichoeres hortulanus	COI	26,500	1491.11	0.08	7.50	115.50	6.46	0.01
(checkerboard wrasse)	cyt *b*	21,600	> 199.50	> 163.50	> 272.50	> 262.50	1.22	0.02
Lutjanus kasmira	COI	41,000	0.23	1282.22	104.50	7.50	0.02	6.14
(bluestripe snapper)	cyt *b*	155,000	15.01	104.61	88.10	3.10	2.82	0.50
Neoniphon sammara	COI	169,000	0.79	0.03	14.85	20.75	0.02	0.001
(Sammara squirrelfish)	cyt *b*	190,000	0.09	0.24	47.48	17.77	0.0025	0.0025
Pygoplites diacanthus	COI	831,000	0.01	0.04	4.20	1.40	0.01	0.01
(regal angelfish)	cyt *b*	> 662,000	0.06	0.04	8.88	11.63	0.0025	0.0025

TITLES AND LEGENDS TO FIGURES

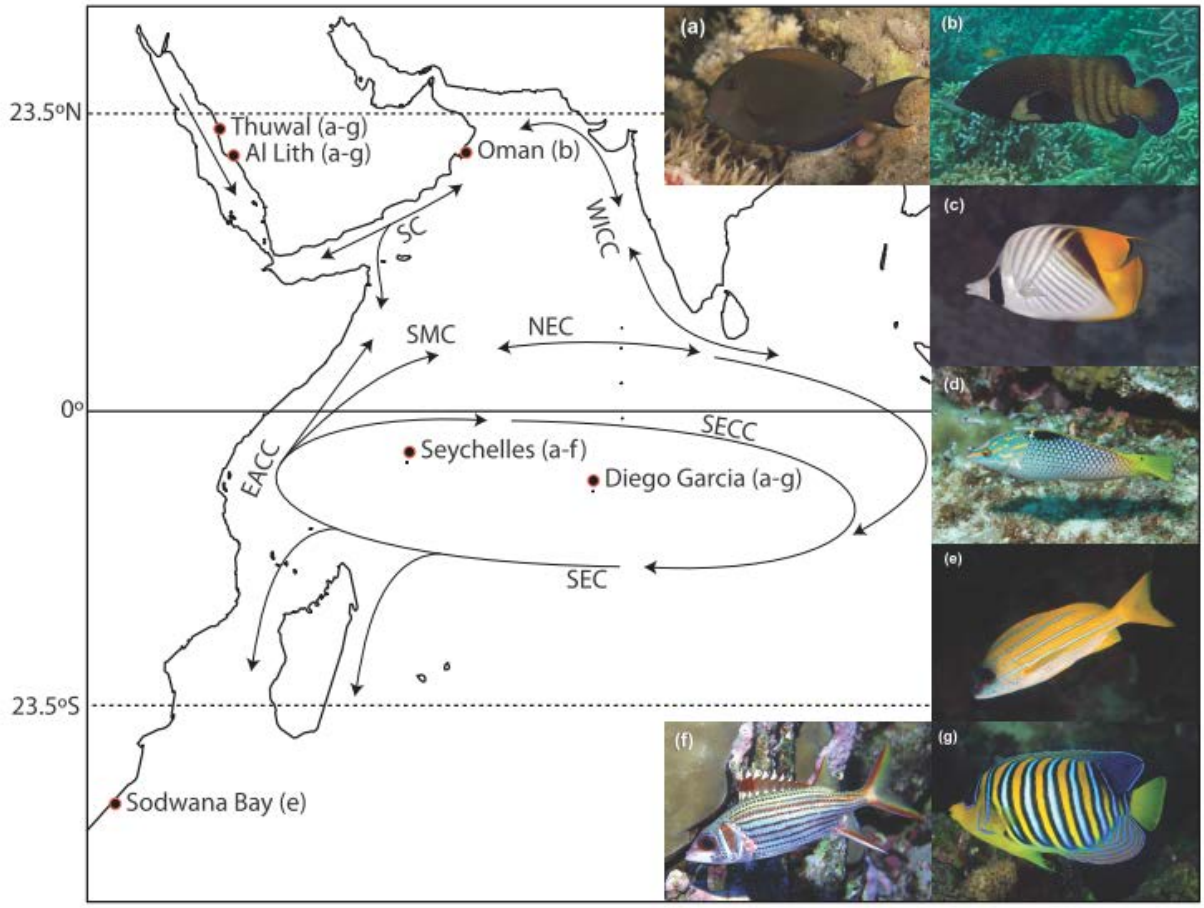
Figure 1 Scaled maps indicating collection sites for all seven reef fish species (a, *Acanthurus nigrofuscus*; b, *Cephalopholis argus*; c, *Chaetodon auriga*; d, *Halichoeres hortulanus*; e, *Lutjanus kasmira*; f, *Neoniphon sammara*; g, *Pygoplites diacanthus*) sampled in the Red Sea and the Western Indian Ocean. Light orange lines show the major current systems flowing through each body of water (abbreviations: EACC, East African Countercurrent; NEC, Northwest Monsoon Current; SC, Somali Current; SEC, South Equatorial Current; SECC, South Equatorial Countercurrent; SMC, Southwest Monsoon Current; WICC, Western Indian Coastal Current). Note the reversing circulation of the SC (from northward to southward), the SMC (from westward to the eastward NEC), the WICC (from eastward to westward), and the current flowing into the Red Sea from the Gulf of Aden (compared with out of the Red Sea and into the Gulf of Aden) during the Northeast Monsoon season (December to March). Site-specific samples sizes are provided in Table 1. (Photo credits: M.L. Berumen, S. Moldzio and L.A. Rocha)

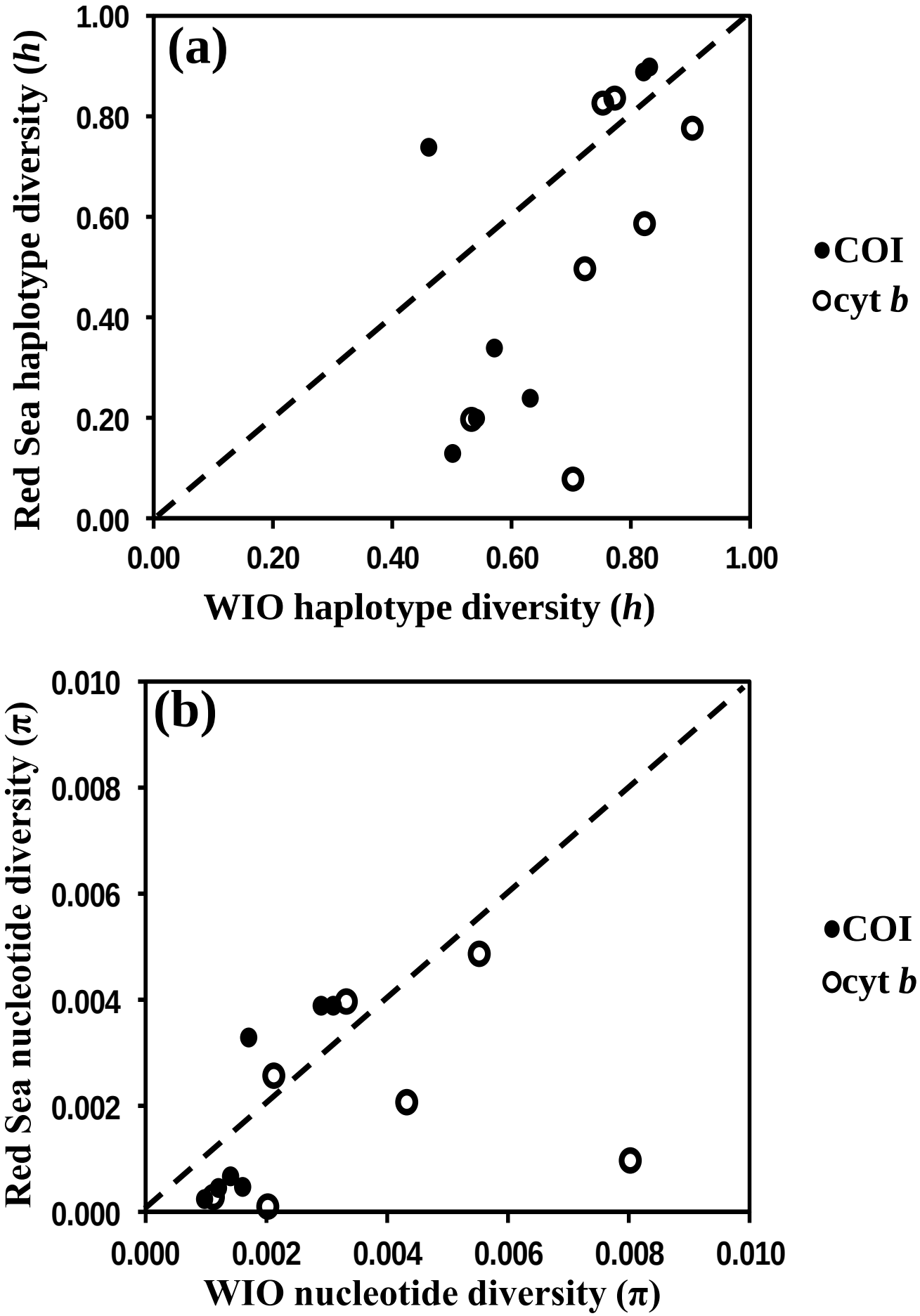
Figure 2 The relationship between (a) haplotype diversity (*h*) or (b) nucleotide diversity (π) estimated for mitochondrial DNA cytochrome-c oxidase subunit-I (COI; black filled circles) and cytochrome *b* (cyt *b*; open circles) genes in the Red

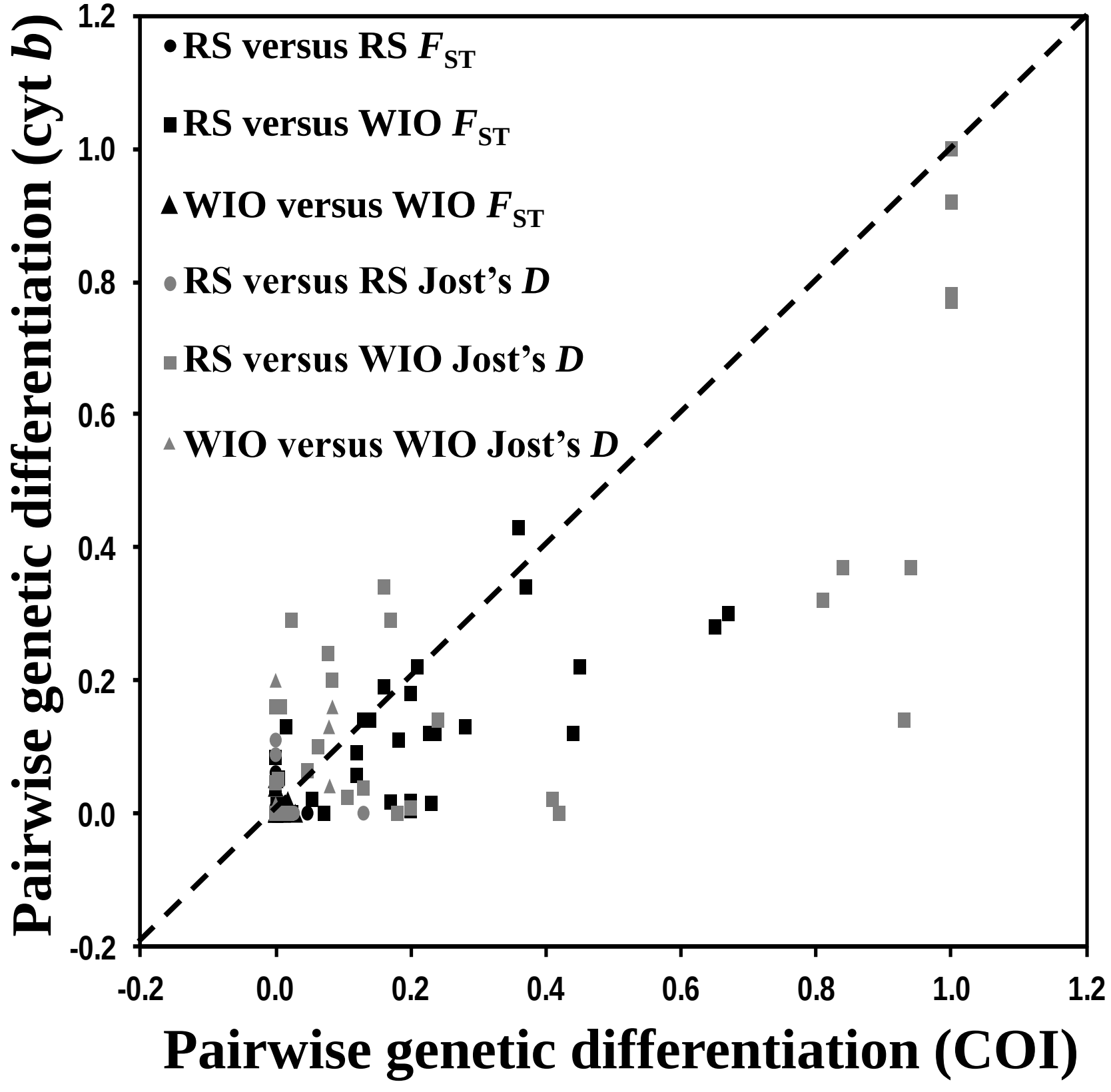
Sea versus the Western Indian Ocean (WIO) populations of species considered in this study. The black dashed line represents a line of unity, which is the point at which genetic diversity estimates in the Red Sea and WIO are equal within a species. Data points above the line of unity indicate greater genetic diversity in the Red Sea, whereas points falling below the line indicate greater genetic diversity in the WIO.

Figure 3 The relationship between mitochondrial DNA cytochrome c oxidase subunit I (COI) and cytochrome *b* (cyt *b*) estimates of pairwise genetic differentiation for the seven species of reef fish, based on comparisons among Red Sea (RS, represented by circles), Western Indian Ocean (WIO; represented by triangles), and Red Sea versus WIO sites (represented by squares). Estimates of both Φ_{ST} (black symbols) and Jost's *D* (grey symbols) are presented here. The black dashed line represents a line of unity, which is the point at which pairwise genetic differentiation estimates between two study sites are equal for each molecular marker. Data points below the line of unity indicate greater genetic differentiation based on COI, whereas points falling above the line indicate greater genetic differentiation based on cyt *b*.

Figure 4 Median-joining networks showing relationships among mitochondrial DNA (a–g) cytochrome-c oxidase subunit-I (COI) and (h–n) cytochrome *b* (cyt *b*) haplotypes for each study species (a/h, *Acanthurus nigrofuscus*; b/i, *Cephalopholis argus*; c/j, *Chaetodon auriga*; d/k, *Halichoeres hortulanus*; e/l, *Lutjanus kasmira*; f/m, *Neoniphon sammara*; g/n, *Pygoplites diacanthus*) collected in the Red Sea (Al Lith and Thuwal) and the Western Indian Ocean (Diego Garcia, Oman, Seychelles, and Sodwana Bay). Each circle represents a unique haplotype and its size is proportional to its total frequency. Branches or black cross-bars represent a single nucleotide change, small black circles represent missing haplotypes, and colours denote collection location as indicated by the embedded key.

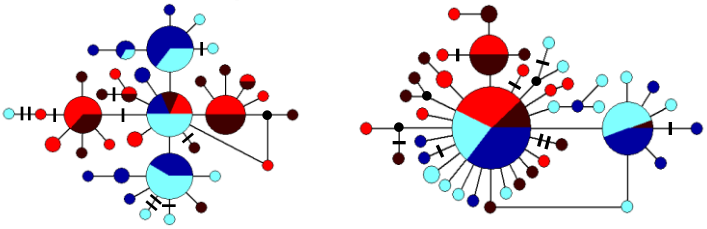

23.5ºN
0º
23.5ºS
Thuwal (a-g)
Al Lith (a-g)
Oman (b)
Seychelles (a-f)
Diego Garcia (a-g)
Sodwana Bay (e)
SC
WICC
SMC
NEC
SECC
EACC
SEC
(a)
(b)
(c)
(d)
(e)
(f)
(g)


(a)
Red Sea haplotype diversity (h)
WIO haplotype diversity (h)
0.00
0.20
0.40
0.60
0.80
1.00
COI
cyt b
(b)
Red Sea nucleotide diversity (π)
WIO nucleotide diversity (π)
0.000
0.002
0.004
0.006
0.008
0.010
COI
cyt b

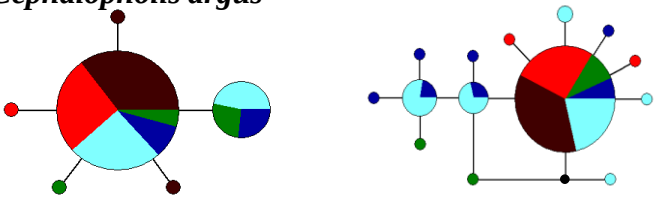

RS versus RS F_{ST}
RS versus WIO F_{ST}
WIO versus WIO F_{ST}
RS versus RS Jost's D
RS versus WIO Jost's D
WIO versus WIO Jost's D
Pairwise genetic differentiation (cyt b)
Pairwise genetic differentiation (COI)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2

COI cyt *b*

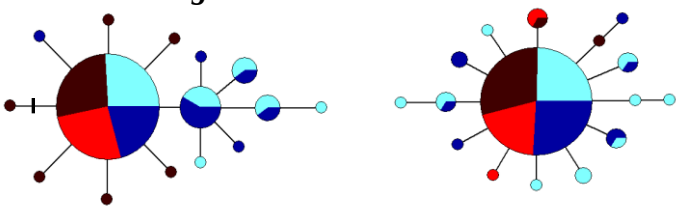
Acanthurus nigrofuscus



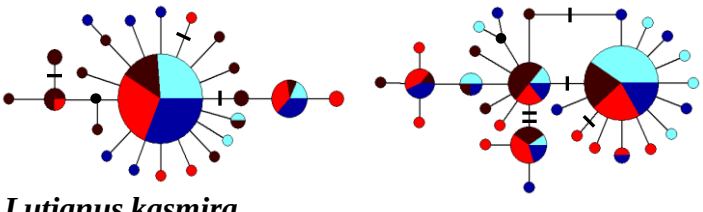
Cephalopholis argus



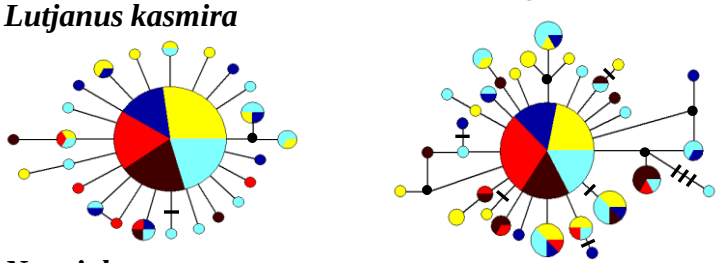
Chaetodon auriga



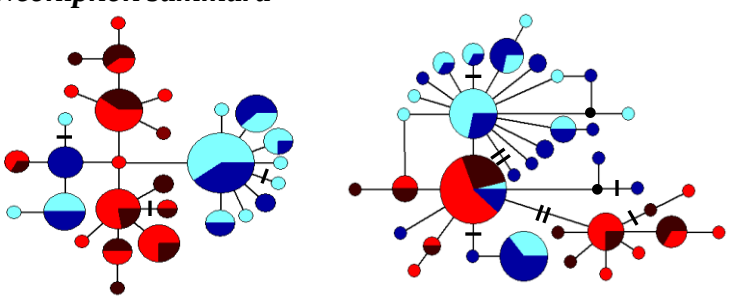
Halichoeres hortulanus



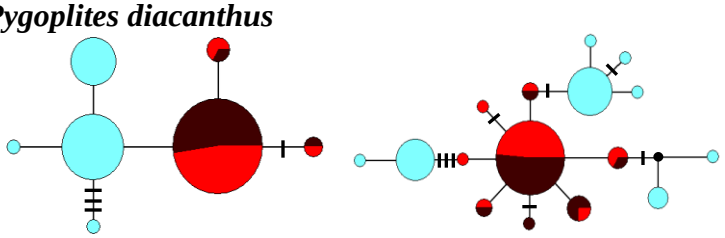
Lutjanus kasmira



Neoniphon sammara



Pygoplites diacanthus



Al Lith
Thuwal
Sodwana Bay
Oman
Seychelles
Diego Garcia